Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Journal of Computational and Applied Mathematics 220 (2008) 548 – 558
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

A new constrained fixed-point algorithm for ordering independent components

Hongjuan Zhang[a], Chonghui Guo[b,*], Zhenwei Shi[c], Enmin Feng[a]

[a] *Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, PR China*
[b] *Institute of Systems Engineering, Dalian University of Technology, Dalian 116024, PR China*
[c] *Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics, Beijing 100083, PR China*

Received 6 June 2007; received in revised form 31 August 2007

Abstract

Independent component analysis (ICA) aims to recover a set of unknown mutually independent components (ICs) from their observed mixtures without knowledge of the mixing coefficients. In the classical ICA model there exists ICs' indeterminacy on permutation and dilation. Constrained ICA is one of methods for solving this problem through introducing constraints into the classical ICA model. In this paper we first present a new constrained ICA model which composed of three parts: a maximum likelihood criterion as an objective function, statistical measures as inequality constraints and the normalization of demixing matrix as equality constraints. Next, we incorporate the new fixed-point (newFP) algorithm into this constrained ICA model to construct a new constrained fixed-point algorithm. Computation simulations on synthesized signals and speech signals demonstrate that this combination both can eliminate ICs' indeterminacy to a certain extent, and can provide better performance. Moreover, comparison results with the existing algorithm verify the efficiency of our new algorithm furthermore, and show that it is more simple to implement than the existing algorithm due to its advantage of not using the learning rate. Finally, this new algorithm is also applied for the real-world fetal ECG data, experiment results further indicate the efficiency of the new constrained fixed-point algorithm.

Keywords: Independent component analysis; Constrained independent component analysis; Lagrange multiplier method; Fixed-point algorithm

1. Introduction

Independent component analysis(ICA) is a very general purpose statistical technique in which observed random data are linearly transformed into components that are maximally independent from each other, and simultaneously have interesting distributions. ICA has become one of the exciting new topics in the field of neural networks, especially unsupervised learning, and more generally in advanced statistics and signal processing. However, most ICA methods suffer from an inherent ambiguity on dilation and permutation. Such indeterminacy cannot be reduced further without additional assumptions [6]. In practice, the ordering of ICs is quite important to separate nonstationary signals or interested signals with significant statistical characters. At present, some methods have been proposed to determine the proper ordering of ICs and most of them belong to two-stage methods. For example, Cheung et al. [4,5] suggested a so-called Testing-and-Acceptance(TnA) algorithm to determine locally optimal ICs ordering. In [18], Wu et al.

* Corresponding author. Tel.: +86 411 84708007.
E-mail address: guochonghui@tshinghua.org.cn (C. Guo).

0377-0427/$ - see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.cam.2007.09.010

considered a more efficient one-pass IC ordering procedure under the mean square error (MSE) criterion. And in [19] a post-procedure step which based on canonical correlation analysis and the prior information of medical signals was added to provide a better ICs ordering. Another ordering technique of constrained ICA was firstly proposed in [11,13]. They introduced constraints into the classical ICA to order the resulted ICs and normalize the demixing matrix in the signal separation procedure. In this approach the separation of ICs and the elimination of indeterminacy are performed simultaneously. Meanwhile it has been shown that this constrained ICA learning algorithm converged approximately three times faster than the two-stage approaches [13]. At present, constrained ICA has been successfully used for some fields such as speech analysis [10], functional MRI data extracting [12–15] and so on. However, a resulting trade off of this method is that the convergence depends on a good choice of the learning rate. A bad choice of the learning rate can, in practice, destroy convergence. Therefore, some ways to make the learning radically fast and reliable may be needed.

In this paper, a simple yet efficient method of constrain ICA is firstly proposed for recovering and ranking the resultant ICs simultaneously. Different from constrained ICA model proposed in [13] which incorporated some statistical measure into the minimizing mutual information criterion, we present a new constrained ICA model which composed of three parts: maximum likelihood criteria as objective function, statistical measure as inequality constraint and the normalization of demixing matrix as equality constraint. Next, we construct a new algorithm which incorporating newFP algorithm into this constrained ICA model. The new algorithm need not be a choice of learning rate which is cumbersome in practical application. It is the new algorithm that overcomes this drawback of the existing learning used in [13]. It is worth mentioning that Lagrange multiplier method is adopted to provide an adaptive solution to this problem. Finally, in order to verify the efficiency of this new algorithm, we analyze its stability of convergence and statistical accuracy, and make comparison with the existing algorithm introduced in [13] for separation of synthesized signals and speech signals. Experiment results show the validity of the proposed algorithm. And comparison results indicate the new algorithm has better performance, and show that it is more simple to implement than the existing algorithm as the new algorithm does not depend on the choice of learning rate. Meanwhile, this new algorithm is applied for the real-world fetal ECG data, these experiment results further demonstrate the efficiency of our new algorithm.

The next section summarizes the classical ICA method and the newFP algorithm. Section 3 introduces the technique of constrained ICA and drives the new constrained fixed-point algorithm. Section 4 demonstrates the present technique with experiments using synthetic signals and speech signals. And in Section 5, we apply the new technique for real-world fetal ECG data. The final section provides discussions and conclusions.

2. Classical ICA

Suppose that there exist M independent source signals $\mathbf{s}(t) = (s_1(t), \ldots, s_M(t))^{\mathrm{T}}$ and N observed mixtures $\mathbf{x}(t) = (x_1(t), \ldots, x_N(t))^{\mathrm{T}}$ of the sources signals (usually $N \geqslant M$). A typical ICA model is

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{1}$$

where $\mathbf{A}$ is an unknown $N \times M$ mixing matrix. The task of classical ICA is to identify an $M \times N$ demixing matrix $\mathbf{W}$ such that the M output signals

$$\mathbf{u}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}\mathbf{A}\mathbf{s}(t) = \mathbf{P}\mathbf{D}\mathbf{s}(t), \tag{2}$$

where $\mathbf{P} \in R^{M\times M}$ is a permutation matrix, $\mathbf{D} \in R^{M\times M}$ is a diagonal scaling matrix, and $\mathbf{u}(t) = (u_1(t), \ldots, u_M(t))^{\mathrm{T}}$. Consequently, the source signals are recovered up to scaling and permutation.

2.1. The Likelihood of the ICA model

From (2), the probability density function of the observations $\mathbf{x}$ can be expressed as [9]:

$$p(\mathbf{x}) = |\det \mathbf{W}|\, p(\mathbf{u}), \tag{3}$$

where $p(\mathbf{u}) = \prod_{i=1}^{M} p_i(u_i)$ is the hypothesized distribution of $p(\mathbf{s})$. Assuming that we have T observations of $\mathbf{x}$, denoted by $\mathbf{x}(1), \mathbf{x}(2), \ldots \mathbf{x}(T)$, then the likelihood can be obtained as the product of this density evaluated at the T points. This is denoted by L and considered as a function of $\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_M)^{\mathrm{T}}$. The log-likelihood is given by [8]:

$$\log L(\mathbf{W}) = \sum_{t=1}^{T} \sum_{i=1}^{M} \log p_i(\mathbf{w}_i^{\mathrm{T}} \mathbf{x}(t)) + T \log |\det(\mathbf{W})| . \tag{4}$$

To simplify notation, we can denote the sum over the sample index t by an expectation operator, and divide the likelihood by T to obtain

$$\frac{1}{T} \log L(\mathbf{W}) = E\left\{\sum_{i=1}^{M} \log p_i(\mathbf{w}_i^{\mathrm{T}} \mathbf{x}(t))\right\} + \log |\det(\mathbf{W})| . \tag{5}$$

The expectation here is not the theoretical expectation, but an average computed from the observed samples. Of course, in the algorithm the expectations are eventually replaced by sample averages, so the distinction is purely theoretical. And maximization of likelihood (5) make the output components independent [8]. A useful preprocessing strategy in ICA is to first whiten the observed vector $\mathbf{x}$. Since for whited data $\widetilde{\mathbf{x}}$ we have $E\{(\mathbf{W}^{\mathrm{T}}\widetilde{\mathbf{x}})^2\} = \mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}$, we can optimize the log-likelihood function on the unit sphere. Therefore, we obtain the following optimization problem:

$$\begin{cases} \max \quad \Psi(\mathbf{W}) = E\left\{\sum\limits_{i=1}^{M} \log p_i(\mathbf{w}_i^{\mathrm{T}} \widetilde{\mathbf{x}}(t))\right\}, \\ s.t. \quad \mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}. \end{cases} \tag{6}$$

2.2. A newFP Algorithm for ICA

We briefly introduce a newFP algorithm for ICA firstly proposed in [17]. Note that at a stable point of the optimization problem (6), the partial derivative of $\Psi(\mathbf{W})$ at $\mathbf{w}_i$ must point in the direction of $\mathbf{w}_i (i = 1, 2, \ldots, M)$, that is, the partial derivative must be equal to $\mathbf{w}_i$ multiplied by some scalar constant. Only in such a case, adding the partial derivative to $\mathbf{w}_i$ does not change its direction and be convergent. This means that we should have

$$\mathbf{W} \leftarrow \frac{\partial \Psi(\mathbf{W})}{\partial \mathbf{W}} = E\{\varphi(\mathbf{u})\widetilde{\mathbf{x}}^{\mathrm{T}}\} = E\{\varphi(\mathbf{u})\mathbf{u}^{\mathrm{T}}\}\mathbf{W}, \tag{7}$$

where $\mathbf{u} = \mathbf{W}\widetilde{\mathbf{x}}(t)$, $\mathbf{W}\mathbf{W}^{\mathrm{T}} = \mathbf{I}$, $\varphi(\mathbf{u}) = (\varphi_1(u_1), \varphi_2(u_2), \ldots, \varphi_M(u_M))^{\mathrm{T}}$; and $\varphi_i(u_i) = (\log p_i(u_i))' = p_i(u_i)'/p_i(u_i) (i = 1, 2, \ldots, M)$. The parametric density estimate $p_i(u_i)$ plays an essential role in the success of the learning rule in Eq.(7). As Lee et al. [9] introduced, the switching between the sub- and super-Gaussian nonlinearities is

$$\varphi_i(u_i) = \begin{cases} -u_i - \tanh(u_i) & \text{super-Gaussian}, \\ -u_i + \tanh(u_i) & \text{sub-Gaussian}. \end{cases}$$

Thus, the learning rule of the newFP algorithm between the sub- and super-Gaussian is

$$\mathbf{W} \leftarrow E\{-\mathbf{K} \tanh(\mathbf{u})\mathbf{u}^{\mathrm{T}} - \mathbf{u}\mathbf{u}^{\mathrm{T}}\}\mathbf{W}, \tag{8}$$

where k_i is the elements of the N-dimensional diagonal matrix $\mathbf{K}$, i.e.,

$$k_i = \mathrm{sign}(E\{\mathrm{sech}^2(u_i)\}E\{u_i^2\} - E\{(\tanh(u_i))u_i\}) = \begin{cases} -1 & \text{super-Gaussian}, \\ 1 & \text{sub-Gaussian}. \end{cases} \tag{9}$$

After every fixed-point iteration, orthogonalization of $\mathbf{W}$ can be done by the symmetric orthogonalization methods [7], i.e.,

$$\mathbf{W} \leftarrow (\mathbf{W}\mathbf{W}^{\mathrm{T}})^{-1/2}\mathbf{W}. \tag{10}$$

3. Constrained ICA

The lack of consistent ordering of components results in different arrangement of ICs each time, which interferes with user to select the available components out of the analysis results. Therefore, an analysis technique that would allow the robust ordering of ICs without the intervention of the user would be rather useful, especially for separating nonstationary signals or signals with significant statistical character. So in the following text, we will introduce how to order ICs through the constrained ICA.

3.1. Ordering and normalizing of ICs

The model of constrained ICA is defined as follows:

$$\begin{cases} \max \quad \Psi(\mathbf{W}) = E\left\{\sum_{i=1}^{M} \log\ p_i(\mathbf{w}_i^{\mathrm{T}}\widetilde{\mathbf{x}}(t))\right\}, \\ s.t. \quad g(\mathbf{W}) \leqslant \mathbf{0}, \quad g(\mathbf{W}) = [g_1(\mathbf{W}), \ldots, g_{M-1}(\mathbf{W})]^{\mathrm{T}}, \\ \qquad h(\mathbf{W}) = \mathbf{0}, \quad h(\mathbf{W}) = [h_1(\mathbf{W}), \ldots, h_M(\mathbf{W})]^{\mathrm{T}}, \end{cases} \tag{11}$$

where $g(\mathbf{W})$ corresponds a set of $M-1$ inequality constraints, $g_i(\mathbf{W}) = I(u_{i+1}) - I(u_i)$ defines the descent order of independent components, where $I(u_i)$ is the index of some statistical measure of the recovered sources, for example, variance $I_{\mathrm{var}} = E\{u_i^2\}$ or normalized kurtosis $I_{\mathrm{kur}} = E\{u_i^4\}/E\{u_i^2\} - 3$. $h(\mathbf{W})$ defines a set of M equality constraints which need $h_i(\mathbf{W}) = \mathbf{w}_i^{\mathrm{T}}\mathbf{w}_i - 1$, $i = 1, \ldots, M$.

Based on the Lagrange multiplier methods, we can firstly define the augmented Lagrange function for inequalities as:

$$\mathscr{L}(\mathbf{W}, \mu) = \Psi(\mathbf{W}) + \frac{1}{2\gamma}\sum_{i=1}^{M-1}\{[\max\{0, \overline{g}_i(\mathbf{W})\}]^2 - \mu_i^2\}, \tag{12}$$

where $\overline{g}_i(\mathbf{W}) = \mu_i + \gamma g_i(\mathbf{W})$. With discrete solution used, the changes of individual w_{ij} can be written as [3]:

$$\begin{aligned} \triangle w_{ij} \propto -\triangle_{w_{ij}}\mathscr{L}(\mathbf{W}(k), \mu(k)) = \max_{w_{ij}} \Psi(\mathbf{W}(k)) \\ - [\max\{0, \overline{g}_{i-1}(\mathbf{W}(k))\} - \max\{0, \overline{g}_i(\mathbf{W}(k))\}]I'_{\mathbf{u}_i}(u_i(k))\widetilde{\mathbf{x}}_j, \end{aligned} \tag{13}$$

where $I'_{u_i}(\cdot)$ is an index measuring first order derivative about u_i, k is the iteration index. The iterative equation for μ_i is

$$\mu_i(k+1) = \max\{0, \mu_i(k) + \gamma[I(u_{i+1}(k)) - I(u_i(k))]\}, \tag{14}$$

where $\gamma > 0$ is penalty parameter. With the learning equation of the newFP algorithm introduced in Section 2, we obtain the following constrained fixed-point algorithm:

$$\mathbf{W} \leftarrow E\{-\mathbf{K}\tanh(\mathbf{u})\mathbf{u}^{\mathrm{T}} - \mathbf{u}\mathbf{u}^{\mathrm{T}}\}\mathbf{W} - \Phi(\mathbf{u})\widetilde{\mathbf{x}}^{\mathrm{T}}, \tag{15}$$

where

$$\Phi(\mathbf{u}) = \begin{bmatrix} -\mu_1 I'(u_1) \\ (\mu_1 - \mu_2)I'(u_2) \\ \vdots \\ (\mu_{M-2} - \mu_{M-1})I'(u_{M-1}) \\ \mu_{M-1}I'(u_M) \end{bmatrix}.$$

Then, it is followed by the normalization of the demixing matrix $\mathbf{W}$. Specifically, the new constrained fixed-point algorithm can be described as follows:

Step 1: Center the observed signals $\mathbf{x}$ to make it mean zero and whiten them to $\widetilde{\mathbf{x}}$.

Step 2: Choose penalty parameter γ and randomly initialize demixing matrix W_0.

Step 3: Choose an initial value for the Lagrange multiplier μ.

Step 4: Update the Lagrange multiplier μ by $\mu(k+1) \leftarrow \max\{0, \mu(k) + \gamma g(\mathbf{W}_k)\}$.
Step 5: Update demixing matrix by $\mathbf{W} \leftarrow E\{-\mathbf{K} \tanh(\mathbf{u})\mathbf{u}^{\mathrm{T}} - \mathbf{u}\mathbf{u}^{\mathrm{T}}\}\mathbf{W} - \Phi(\mathbf{u})\widetilde{\mathbf{x}}^{\mathrm{T}}$.
Step 6: Normalize $\mathbf{W}$ by $\mathbf{W} \leftarrow (\mathbf{W}\mathbf{W}^{\mathrm{T}})^{-1/2}\mathbf{W}$.
Step 7: If not converged, go back to Step 4.

4. Computer simulations and performance analyses

Notice that we use normalized kurtosis as measure to order ICs in the following experiments. We demonstrate our algorithm with experiments using synthetic data and speech data. In addition we compare the new algorithm with the existing algorithm in [13]. The existing algorithm is:

$$\mathbf{W} \leftarrow \mathbf{W} + \eta(\mathbf{I} + (\varphi(\mathbf{u}) + \Phi(\mathbf{u}))\mathbf{u}^{\mathrm{T}})\mathbf{W}, \tag{16}$$

where $\varphi(\mathbf{u})$, $\Phi(\mathbf{u})$ are the same as above, η is the learning rate. For convenience, we will denote (16) as Algorithm 1.

The accuracy of the recovered ICs compared to the sources are expressed using the signal-to-noise ratio (SNR) in dB given by:

$$\mathrm{SNR} = 10\log_{10}(s^2/\mathrm{MSE}), \tag{17}$$

where s^2 denotes the variance of the source and MSE denotes the mean square error between the original signal and recovered signal. The higher SNR is, the better the performance is. Besides, the performance of algorithms is measured using the performance index (PI), defined as [1]:

$$\mathrm{PI} = \sum_{i=1}^{M}\left(\sum_{j=1}^{M}\frac{|p_{ij}|}{\max_k |p_{ik}|} - 1\right) + \sum_{j=1}^{M}\left(\sum_{i=1}^{M}\frac{|p_{ij}|}{\max_k |p_{kj}|} - 1\right), \tag{18}$$

where p_{ij} is the *ij*th element of $M \times M$ matrix $\mathbf{P} = \mathbf{W}\mathbf{A}$. The larger the value PI is, the poorer the statistical performance of a separation algorithm is. Note that all the experiments are repeated over 100 different realizations of the input data. For each of the 100 realizations, each algorithm is run N times (here $N = 200$) with different mixing matrices, and for every trial, the error is estimated by the average of the errors.

4.1. Experiments with synthesized signals

Five random signals (500 data points): one Gaussian, two sub-Gaussian and two super-Gaussian were used for simulations (available at http://www.cis.hut.fi/projects/ica). Notice that in Algorithm 1, the learning rate η must be

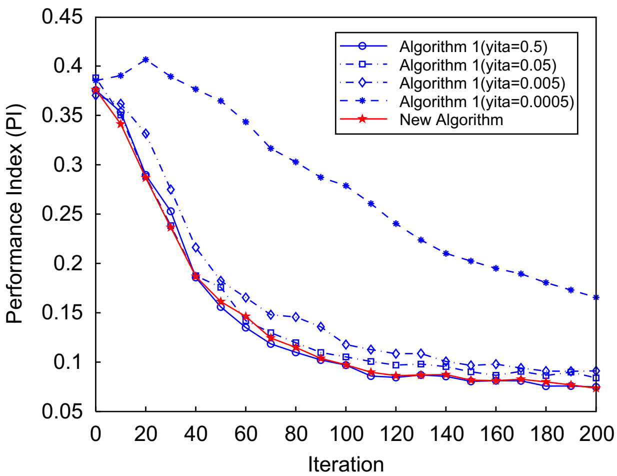

Fig. 1. Convergence of two algorithms for synthesized signals.

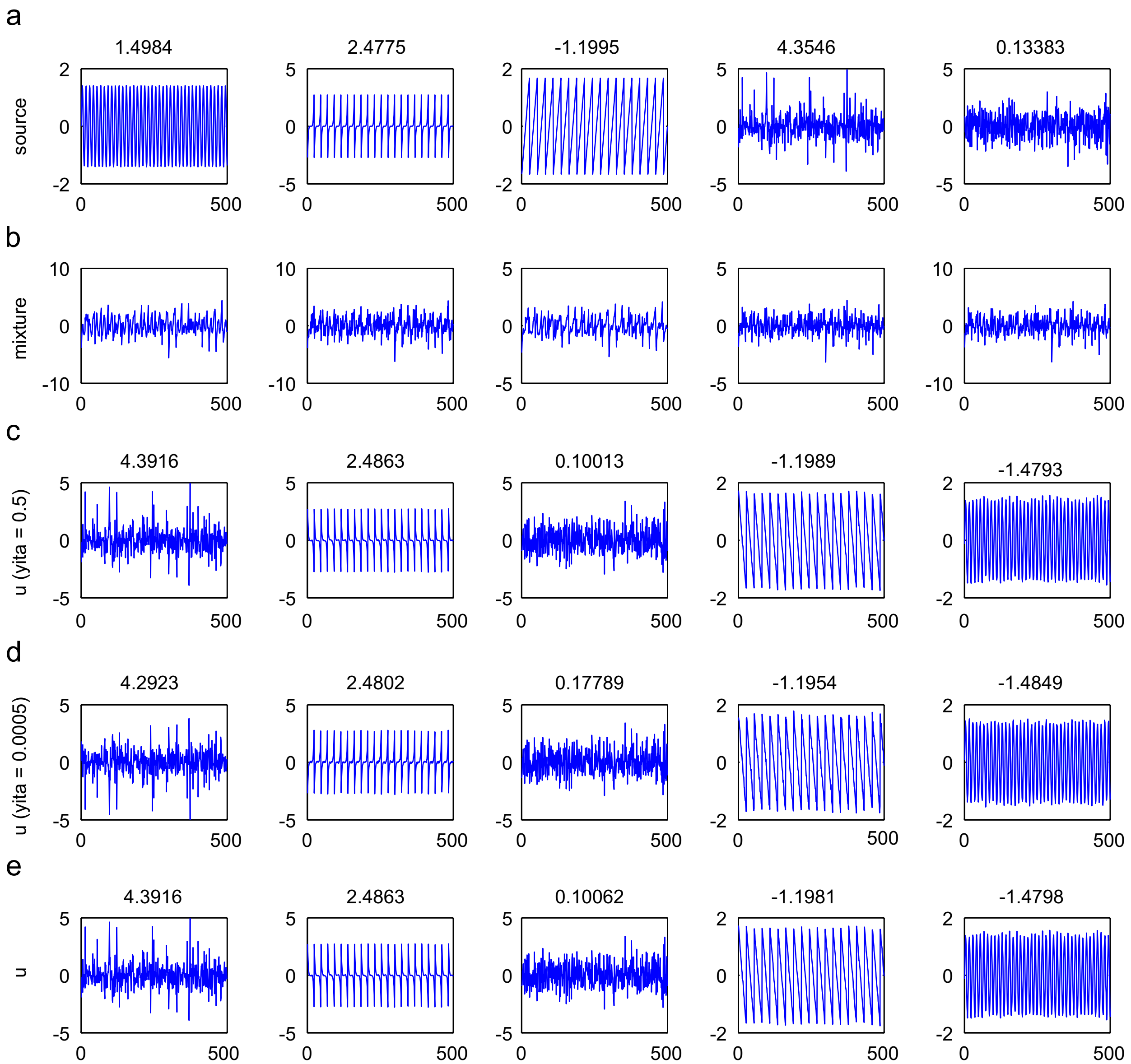

Fig. 2. (a) Sources of synthesized signals; (b) Mixtures of synthesized signals; (c) Recovered signals of Algorithm 1 ($\eta = 0.5$); (d) Recovered signals of Algorithm 1 ($\eta = 0.0005$); (e) Recovered signals of the new algorithm.

chosen at the beginning. So we firstly chose the proper value of parameter η in this algorithm. Fig. 1 shows the convergence of Algorithm 1 and the new algorithm. When η equals 0.5, the average PI value is the lowest while $\eta = 0.0005$ corresponding to the highest one. Meanwhile we can also see that the new algorithm is better than all of Algorithm 1. In order to verify this conclusion, we keep on doing the following separation experiments. The waveforms of five original signals, their mixtures, and the separated output signals are shown in Fig. 2, where Fig. 2(c)–(e) show the signals recovered by Algorithm 1 (when $\eta = 0.5$ and 0.0005) and the new algorithm, respectively. The normalized kurtosis values of original and recovered signals are labeled in the Fig. 2, and the recovered signals of two algorithms are placed in a descending order. For exactly comparing their performance in quantity, more precise comparisons about SNR index are also shown in Table 1.

Table 1
The SNRs(dB) of output components using two algorithms for synthesized signals

Output	Algorithm 1 ($\eta = 0.0005$)	Algorithm 1 ($\eta = 0.5$)	New algorithm
u1	20.21	24.94	25.01
u2	26.60	31.21	31.27
u3	15.48	20.21	20.40
u4	23.22	28.67	28.32
u5	22.00	22.14	22.28

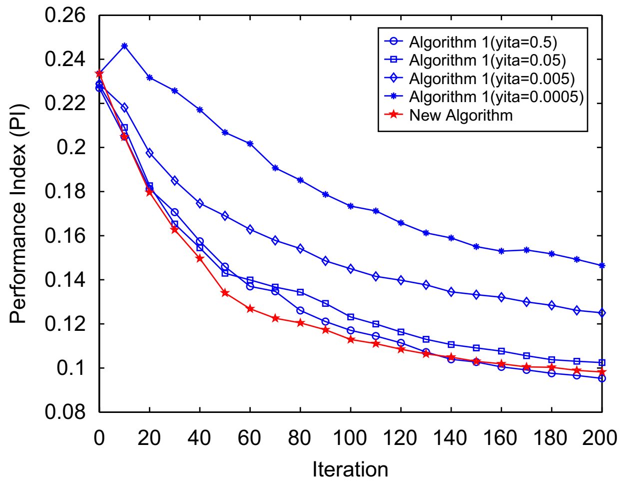

Fig. 3. Convergence of two algorithms for speech signals.

It can be seen that the SNR index of the new algorithm are almost the best among them. As to Algorithm 1, its SNR index has a bigger increase from $\eta = 0.0005$ to 0.5. So Algorithm 1 is significantly influenced by the value of the learning rate, and only when taking the best one, the experiment results are able to achieve a better performance just like the new algorithm.

4.2. *Experiments with speech signals*

The new constrained fixed-point algorithm is also be used to extract speech signals from their mixtures in an order. Four speech signals (500 data points): one Gaussian, two sub-Gaussian and one super-Gaussian were used for simulations. Similar with the above experiment, we firstly chose the proper value of parameter η in Algorithm 1. Fig. 3 presents the performance of two algorithms. From Fig. 3, we can see that in this experiment the best learning rate of Algorithm 1 equals 0.5 and the worst is 0.0005 (which is a coincidence with the experiment of above experiment). And these comparisons indicate only when we select the learning rate properly, Algorithm 1 is able to have a good performance similar to the new algorithm which need not choose the learning rate. This result can be confirmed by Fig. 4 further which presents the output signals recovered by Algorithm 1 ($\eta = 0.5$ and 0.0005, respectively) and the new algorithm. Meanwhile, the SNR values of the corresponding recovered signals are shown in Table 2. It can be found that the new algorithm recovers signals in the highest SNR. However, in Algorithm 1 when η equals 0.0005, the values of SNR are the lowest than Algorithm 1 ($\eta = 0.5$) and the new algorithm.

Based on above experiment results, we can see the new algorithm is superior to Algorithm 1 from two aspects: the PI index of the new algorithm is smaller and the SNR index is higher. So experiment results verify the effectiveness of new algorithm. Meanwhile from these experiments we can know that the performance of Algorithm 1 excessively relies on the selection of the learning rate, while the new constrained fixed-point algorithm overcomes this disadvantage perfectly.

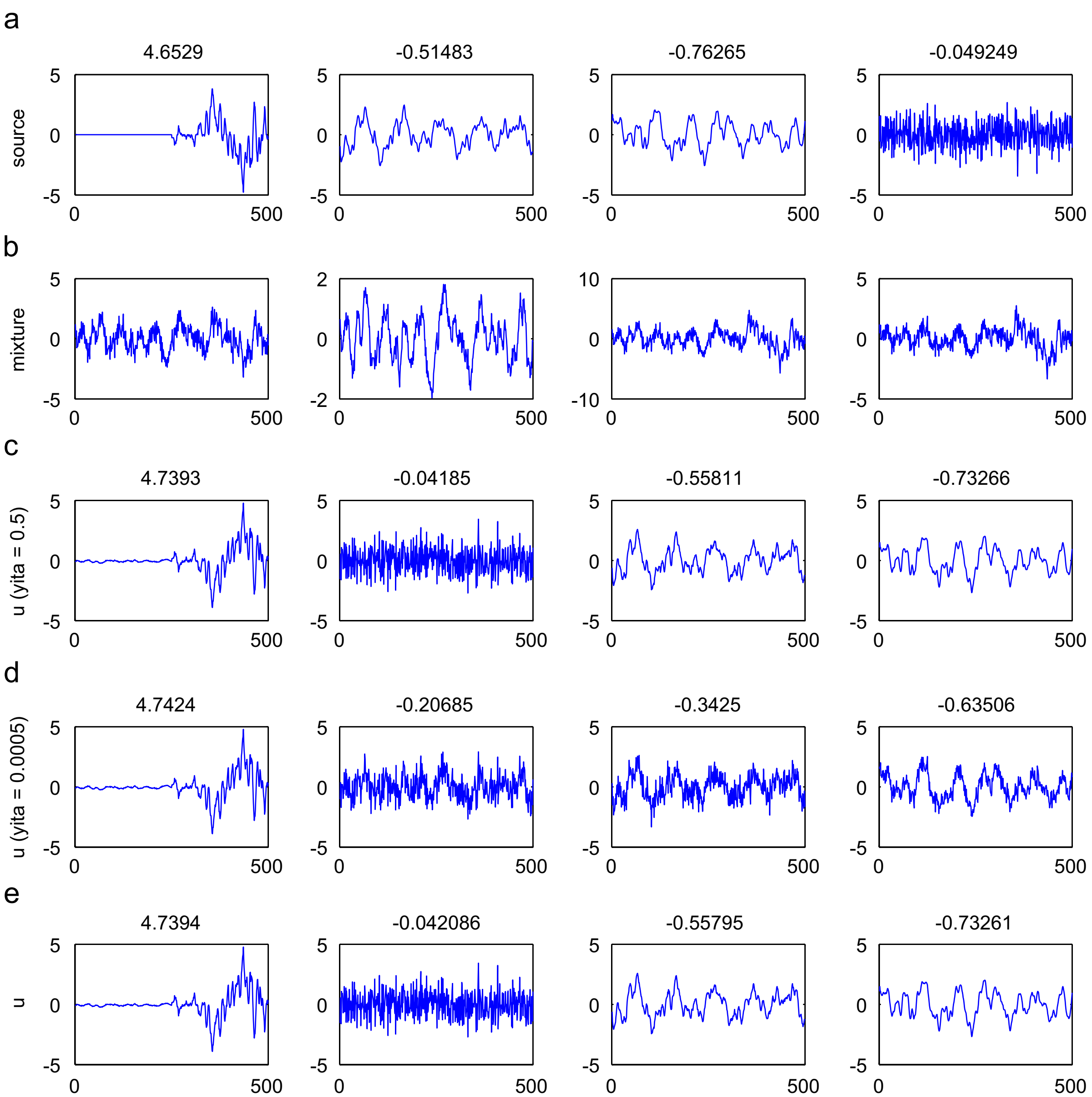

Fig. 4. (a) Sources of speech signals; (b) Mixtures of speech signals; (c) Recovered signals of Algorithm 1 ($\eta = 0.5$); (d) Recovered signals of Algorithm 1 ($\eta = 0.0005$); (e) Recovered signals of the new algorithm.

Table 2
The SNRs(dB) of output components using two algorithms for speech signals

Output	Algorithm 1 ($\eta = 0.0005$)	Algorithm 1 ($\eta = 0.5$)	New algorithm
u1	22.08	23.33	23.33
u2	3.13	29.61	29.98
u3	3.76	12.23	12.24
u4	8.27	19.09	19.10

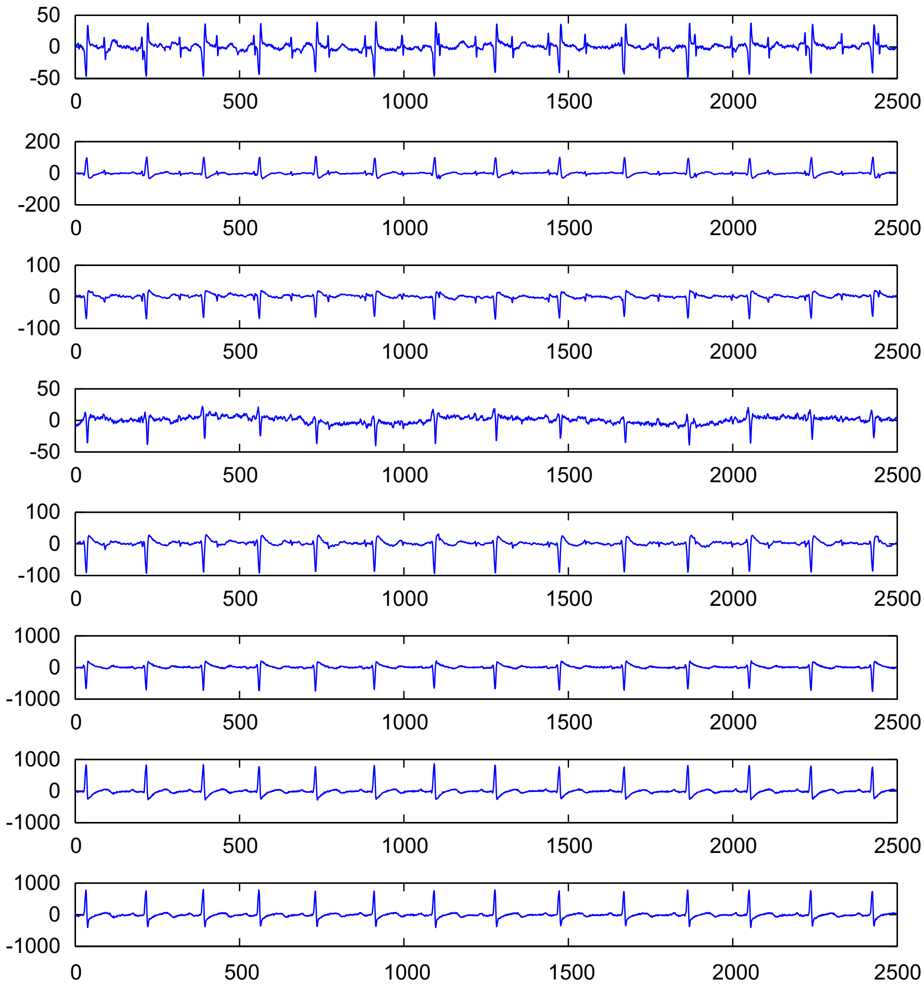

Fig. 5. The 8-channel of ECG recording obtained from a pregnant woman.

5. Experiments with real-world data

To check the practicality of the proposed algorithm, we have performed experiment on real-world ECG data which was distributed in [16]. This data is a famous electrocardiogram measured from a pregnant woman (in Fig. 5). One can see the heart beating of both the mother (stronger and slower) and the fetus (weaker and faster). Note that the fetal influence is stronger in the first channel in Fig. 5. The ECG measurements are recorded over 10 s and sampled at 250 Hz (although in De Moor's homepage he claims the sampling frequency is 500 Hz, Barros et al. [2] assure it is 250 Hz). The task is to obtain source signals orderly which included fetal ECG, mother ECG, mother's respiration and some noises due to the electronic equipments. Since the mixing process and the pure source signals are unavailable not like above simulation experiments, the PI performance cannot be computed as above. So it is not easy to present the suitable learning rate in Algorithm 1. However, the new constrained fixed-pointed algorithm can avoid this problem. Therefore this extracting experiment can be made only by the new algorithm. The waveforms of the extracted signals are shown in Fig. 6, where Fig. 6(a) and (b) give us mother ECG, simultaneously fetal ECG is shown in Fig. 6(c) clearly [2], and Fig. 6(d–h) should be mother's respiration signals and other noise signals. It is clear to see that the extracted signals are correct and are ranked according to their kurtosis values in a descending order according to the kurtosis values. The experiment with ECG data further demonstrate the availability of our new algorithm in separating and ordering ICs simultaneously in real-world data.

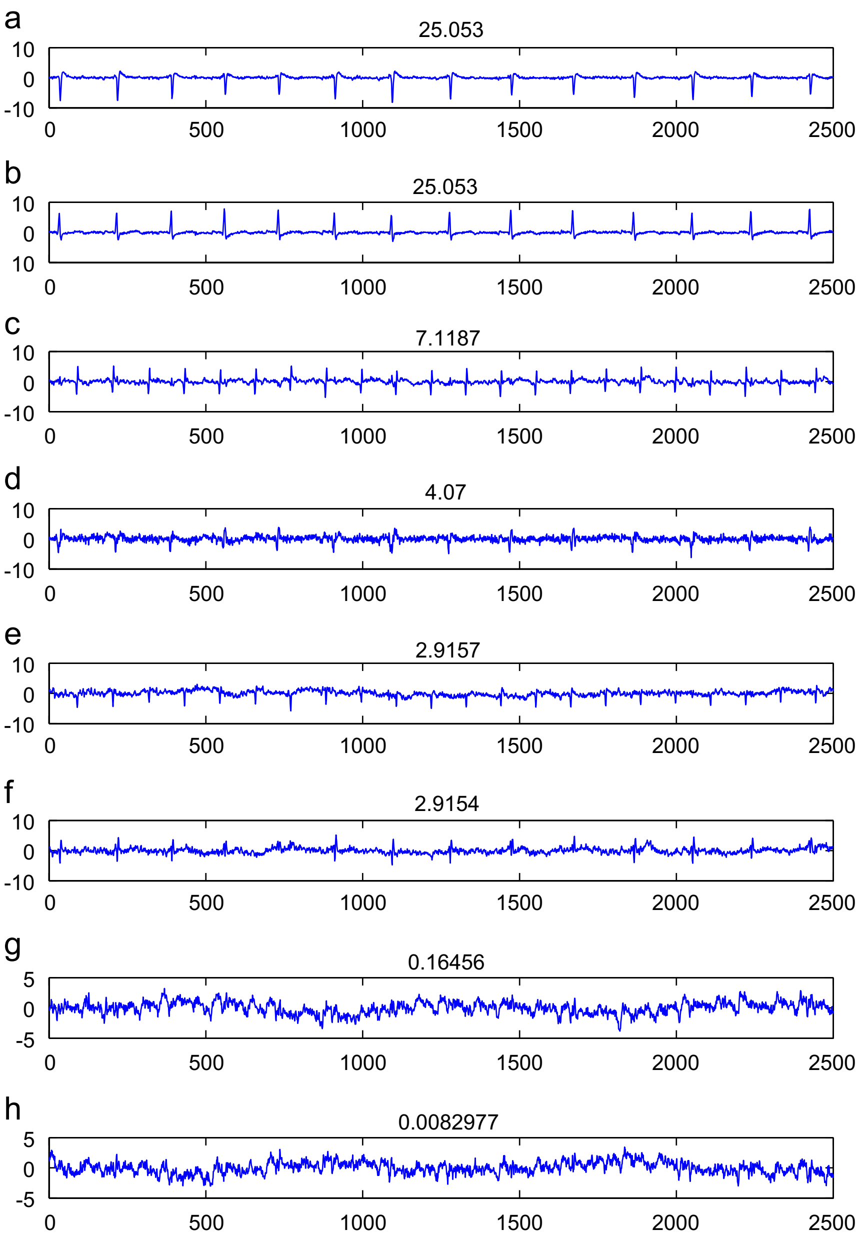

Fig. 6. The recovered ECG signals using new algorithm.

6. Discussions and conclusions

In this work, we present a simple yet efficient algorithm for ranking the resulting components of the ICA algorithm. This technique is achieved through (1) adding maximum likelihood criterion as objective function, statistical measure as inequality constraint and the normalization of demixing matrix as equality constraint; (2) combining the newFP algorithm with this constrained ICA model to construct a new constrained fixed-point algorithm. This combination both can eliminate ICs' indeterminacy on permutation and dilation to a certain extent, and can provide better performance for recovering synthesized signals and speech signals. Moreover, we make comparison with the existing algorithm in [13], comparison results verify the efficiency of our new algorithm furthermore. According to above experiment results, we must point out that the existing algorithm is significantly influenced by the value of the learning rate, and only when taking the best one, the experiment results are able to achieve a better performance like our new algorithm. Meanwhile we also applied this new algorithm to real-world fetal ECG data to demonstrate its efficiency further. Therefore the new constrained fixed-point algorithm is a more simple and practical technique for recovering and ranking the resultant ICs simultaneously.

Acknowledgments

This work is supported by Natural Science Foundation of China under Grant no. 10571018, 60605002. The authors would like to thank the referees and the editorial board for their insightful comments and suggestions.

References

[1] S.I. Amari, A. Cichocki, H. Yang, A new learning algorithm for blind source separation, in: Advances in Neural Information Processing Systems, vol. 8, 1996, pp. 757–763.
[2] A.K. Barros, A. Cichocki, Extraction of specific signals with temporal structure, Neural Comput. 13 (9) (2001) 1995–2003.
[3] D.P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods, Academic Press, New York, 1982.
[4] Y. Cheung, L. Xu, An expirical method to selected dominant independent components in ICA for time series analysis, in: Proceeding of 1999 International Joint Conference on Neural Networks, vol. 6, 1999, pp. 3883–3887.
[5] Y. Cheung, L. Xu, Independent component ordering in ICA series analysis, Neurocomputing 41 (1–4) (2001) 145–152.
[6] P. Comon, Independent component analysis: A new concept?, Signal Process. 36 (1994) 287–314.
[7] A. Hyvärinen, Fast and robust fixed-point algorithm for independent component analysis, IEEE Trans. Neural Networks 10 (3) (1999) 626–634.
[8] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, Wiley, New York, 2001.
[9] T.W. Lee, M. Girolami, T. Sejnowski, Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources, Neural Comput. 11 (2) (1999) 417–441.
[10] Q.H. Lin, Y.R. Zheng, F.L. Yin, H.L. Liang, Speech segregation using constrained ICA, Lecture Notes in Comput. Sci. 3173 (2004) 755–760.
[11] W. Lu, J.C. Rajapakse, Constrained independent component analysis, in: Advances in Neural Information Processing systems, vol. 13, 2000, pp. 570–576.
[12] W. Lu, J. C. Rajapakse, ICA with reference, in: Proceedings of the 3rd International Conference on Independent Component Analysis and Blind Source Separation(ICA2001), 2001, pp. 120–125.
[13] W. Lu, J.C. Rajapakse, Eliminating indeterminacy in ICA, Neurocomputing 50 (2003) 271–290.
[14] W. Lu, J.C. Rajapakse, Approach and applications of constrained ICA, IEEE Trans. Neural Networks 16 (1) (2005) 203–212.
[15] W. Lu, J.C. Rajapakse, ICA with reference, Neurocomputing 69 (2006) 2244–2257.
[16] D. De Moor (Ed.), Daisy: Database for the Identification of Systems, 1997 available online at: ⟨http://www.esat.kuleuven.ac.be/sista/daisy⟩.
[17] Z. Shi, H. Tang, Y. Tang, A new fixed-point algorithm for independent component analysis, Neurocomputing 56 (2004) 467–473.
[18] H.C. Wu, L.H. Yu, W.K. Li, An independent component ordering and selection procedure based on the MSE criterion, Lecture Notes in Comput. Sci. 3889 (2006) 286–294.
[19] T. Youssef, A.-B.M. Youssef, S. LaConte, X.P. Hu, Y.M. Kadaha, Robust ordering of independent components in functional magnetic resonance imaging time series data using canonical correlation analysis, in: Proceedings of SPIE 5031, 2003, pp. 332–340.